Популярные лекции
ПО МАТЕМАТИКЕ

И. Р. ШАФАРЕВИЧ

О РЕШЕНИИ УРАВНЕНИЙ ВЫСШИХ СТЕПЕНЕЙ

(МЕТОД ШТУРМА)

*

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ТЕХНИКО-ТЕОРЕТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА · 1954

ОГЛАВЛЕНИЕ

Введение 3

§ 1. Границы корней 4

§ 2. Общие корни многочленов и равные корни 7

§ 3. Характеристнка пары многочленов 9

§ 4. Число корней многочлена, лежащих между a и b 17

И. Р. Шафаревич. О решении уравнений высших степеней.

Редактор *В. Б. Орлов.*

Техн. редактор *С. С. Гаврилов.* Корректор *Ц. С. Варшавская*

Сдано в набор 25/V 1954 г. Подписано к печати 22/VII 1954 г. Бумага 84×108¹/₃₂. Физ. печ. л. 0,75. Условн. печ. л. 1,23. Уч.-изд. л. 1,24. Тираж 30 000 экз. Т-04880. Цена книги 40 к. Заказ № 1465.

Государственное издательство технико-теоретической литературы. Москва, Б. Калужская, 15.

4-я типография им. Евг. Соколовой Союзполиграфпрома Главиздата Министерства культуры СССР. Ленинград, Измайловский пр., 29.

ВВЕДЕНИЕ

В курсе алгебры средней школы выводится формула для решения квадратного уравнения, а из курса физики видно, насколько необходима эта формула для решения многих физических вопросов (например, в задачах, связанных с равноускоренным движением, и т. д.).

Не меньшую роль, чем квадратные уравнения, играют в математике и ее приложениях уравнения третьей и более высоких степеней. Люди почти так же давно начали заниматься уравнениями высших степеней, как и квадратными уравнениями. Известны вавилонские клинописные таблички, в которых решаются некоторые кубические уравнения. Несмотря на то, что этим вопросом занимались так давно, основные факты об уравнениях высших степеней были открыты только в XIX веке. Эта лекция посвящена обзору некоторых основных свойств уравнений высших степеней.

Способ, которым мы будем выводить свойства уравнений высших степеней, резко отличается от того способа, при помощи которого в курсе алгебры средней школы выводят свойства квадратных уравнений. Почти все свойства квадратных уравнений выводятся из формулы для их решения, мы же не будем выводить формулу для решения уравнений высших степеней, а получим их свойства из некоторых общих алгебраических и геометрических соображений.

Дело в том, что для большинства уравнений высших степеней не существует такой формулы, как для уравнений второй степени. В тех же случаях, где такая формула есть, она настолько сложна, что из нее невозможно вывести никаких свойств уравнения. Но и независимо от этого, наш путь имеет еще одно преимущество: он делает более ясной истинную причину тех фактов, которые доказываются.

Все рассуждения, которые здесь будут приведены, годятся для уравнений *любой* степени. Часто они будут изложены в общем виде. В некоторых же случаях, когда рассуждение в общем случае принципиально то же, но удлиняет выкладку, мы будем приводить его лишь для уравнений третьей степени и только формулировать то, что получится в общем случае. Очень рекомендуется провести все рассуждения самостоятельно в общем случае.

Наконец, совсем выпущены доказательства фактов, подобных следующему: если график многочлена имеет точки по разные стороны оси x, то он эту ось пересекает. Вероятно, некоторые читатели не почувствуют потребности в доказательстве подобных предложений. Тот же, кто пожелает провести эти доказательства, легко сделает это при помощи простейших свойств непрерывных функций, которые можно узнать из первых глав любого курса анализа.

В этой книжке мы будем заниматься только свойствами действительных корней уравнений, так что от читателя не потребуется знания свойств комплексных чисел. Заметим, что свойства комплексных корней уравнений могут быть выведены с помощью таких же методов, но несколько усложненных.

§ 1. ГРАНИЦЫ КОРНЕЙ

Первая задача, которую мы себе поставим, заключается в следующем: *для каждого уравнения указать границы, между которыми расположены его корни.*

Предположим, что наше уравнение — третьей степени и имеет вид

$$ax^3 + bx^2 + cx + d = 0. \qquad (1)$$

Мы покажем сейчас, как найти такое положительное число N, что если абсолютная величина x превосходит N, то левая часть уравнения отлична от нуля. Тогда корни будут наверняка расположены между $-N$ и N. Для этого постараемся выбрать N таким, что если абсолютная величина x больше N, то абсолютная величина первого члена превосходит абсолютную величину суммы трех остальных. Тогда первый член не сможет сократиться с суммой остальных и все выражение будет отлично от нуля.

Мы заведомо достигнем своей цели, если треть абсолютной величины первого члена будет больше абсолютной вели-

чины каждого из трех остальных, т. е.

$$\frac{1}{3}|a||x|^3>|b||x|^2,\quad \frac{1}{3}|a||x|^3>|c||x|,$$

$$\frac{1}{3}|a||x|^3>|d|.$$

Решая эти неравенства, получим:

$$|x|>3\frac{|b|}{|a|},\quad |x|>\sqrt{3\frac{|c|}{|a|}},\quad |x|>\sqrt[3]{3\frac{|d|}{|a|}}.$$

Таким образом, мы получим число с нужным нам свойством, взяв за N большее из трех чисел $3\frac{|b|}{|a|}$, $\sqrt{3\frac{|c|}{|a|}}$, $\sqrt[3]{3\frac{|d|}{|a|}}$. В самом деле, при $|x|$, большем N, все три неравенства будут выполняться, а следовательно, левая часть уравнения (1) не будет обращаться в нуль.

Для уравнения степени n,

$$ax^n+bx^{n-1}+cx^{n-2}+\ldots+kx+l=0, \qquad (2)$$

надо было бы принять за N наибольшее из ряда чисел:

$$n\frac{|b|}{|a|},\quad \sqrt{n\frac{|c|}{|a|}},\quad \sqrt[3]{n\frac{|d|}{|a|}},\quad \ldots,\quad \sqrt[n-1]{n\frac{|k|}{|a|}},\quad \sqrt[n]{n\frac{|l|}{|a|}}.$$

Заметим, что мы доказали несколько больше, чем утверждали. Так как абсолютная величина первого члена в левой части уравнения (1) больше, чем абсолютная величина суммы остальных членов, то знак всего выражения определяется знаком первого члена. Таким образом, мы не только знаем, что левая часть уравнения (1) есть число, отличное от нуля, когда абсолютная величина x превосходит N, но можем указать и знак этого числа — он совпадает со знаком первого члена.

Уже из этих простых соображений мы можем вывести важные следствия о корнях уравнений. Для этого надо воспользоваться графиком функции

$$y=ax^n+bx^{n-1}+\ldots+kx+l. \qquad (3)$$

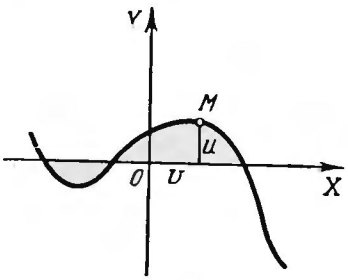
Пусть (черт. 1) на плоскости взяты оси координат и вычерчен график функции (3). Согласно общему правилу черчения гра-

фиков ордината u точки M графика равна тому числу, которое получится, если подставить ее абсциссу v в выражение (3) вместо x. В частности, ордината обратится в нуль, когда абсцисса будет корнем уравнения (2). Это значит, что корни уравнения геометрически изображаются точками пересечения графика с осью x.

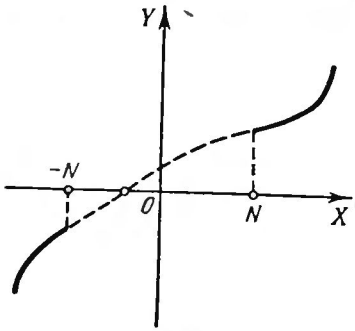
Пусть наше уравнение — третьей степени. Мы разделим обе части его на коэффициент a при первом члене и запишем в виде

$$x^3 + px^2 + qx + r = 0. \qquad (4)$$

Посмотрим, что̀ геометрически обозначает найденное число N с указанными выше свойствами. Все корни уравне-



Черт. 1.



Черт. 2.

ния (4) расположены между $-N$ и N. Это значит, что график функции

$$y = x^3 + px^2 + qx + r \qquad (5)$$

может пересекать ось x только в точках, абсциссы которых лежат между $-N$ и N. Но обязан ли график пересекать ось x? Вспомним, что если x по абсолютной величине больше N, то знак функции (5) совпадает со знаком первого члена. Но знак первого члена нам известен — он совпадает со знаком x. Таким образом, если x больше N, то многочлен положителен, а это значит, что график в этой части лежит выше оси x. Если же x меньше $-N$, то многочлен отрицателен, т. е. график в этой части лежит ниже оси x. То, что мы таким образом узнаем о графике, нарисовано на черт. 2. Из этого чертежа ясно, что график должен пере-

сечь хоть один раз ось x, а это значит, что уравнение третьей степени имеет хотя бы один корень. Аналогичный факт имеет место и для уравнений высших степеней: уравнение нечетной степени имеет хотя бы один корень. Ясно, что для уравнений четной степени эта теорема не имеет места, как показывают уже квадратные уравнения, которые могут вовсе не иметь корней *).

§ 2. ОБЩИЕ КОРНИ МНОГОЧЛЕНОВ И РАВНЫЕ КОРНИ

Основным алгебраическим приемом, которым мы будем дальше пользоваться, является деление многочлена на многочлен с остатком. Если имеется два многочлена f и g, то мы можем разделить (обычным способом, «уголком») тот из них, который имеет бо́льшую степень, на другой, и получим неполное частное q и остаток h, который имеет уже меньшую степень, чем делитель. Это можно записать в виде формулы

$$f = g \cdot q + h. \tag{6}$$

Деление с остатком — это общий прием, который сводит изучение пары многочленов f и g к изучению пары многочленов g и h, имеющих меньшие степени. Именно благодаря этому обстоятельству часто оказывается, что свойства пары многочленов могут быть легче изучены, чем свойства одного многочлена.

В качестве примера рассмотрим, как находятся общие корни двух многочленов f и g. Если α есть корень как многочлена f, так и многочлена g, то, полагая в соотношении (6) $x = \alpha$, мы получим, что и h должно обратиться в нуль, т. е. иметь корнем α. Таким образом, общий корень многочленов f и g является общим корнем многочленов g и h. Наоборот, из того же соотношения (6) мы видим, что если α есть общий корень многочленов g и h, то он является также корнем f, т. е. общим корнем многочленов f и g. Взятые вместе, эти два утверждения показывают, что общие корни многочленов f и g, с одной стороны, и многочленов g и h, с другой, совпадают. Достижение здесь заключается в том, что многочлены g и h имеют меньшие степени, чем f и g.

*) Напомним, что мы под корнем всегда подразумеваем вещественный корень, так что, например, уравнение $x^2 + 1 = 0$, с нашей точки зрения, не имеет корней.

С g и h мы можем повторить те же рассуждения, разделив g с остатком на h. Таким образом, мы будем получать пары многочленов все меньших степеней, причем общие корни у всех этих пар многочленов будут одни и те же. Мы при нуждены будем остановиться, когда придем к паре многочленов u и v, один из которых, например v, есть нуль. Но в этом случае все корни многочлена u являются общими корнями u и v, так как v, будучи равно нулю, имеет все числа своими корнями.

Таким образом, нахождение общих корней многочленов f и g свелось к нахождению корней многочлена u, имеющего, как правило, гораздо меньшую степень.

Рассмотрим пример. Найти общие корни многочленов x^4+x^2+3x+1 и x^3+x+2.

Делим первый с остатком на второй:

$$\begin{array}{l|l} x^4+x^2+3x+1 & x^3+x+2 \\ \underline{x^4+x^2+2x} & \underline{} \\ x+1 & x \end{array}$$

Остаток есть $x+1$. Следовательно, общие корни у первоначальной пары многочленов те же, что и у многочленов x^3+x+2 и $x+1$.

Делим опять первый с остатком на второй:

$$\begin{array}{l|l} x^3+x+2 & x+1 \\ \underline{x^3+x^2} & \underline{} \\ -x^2+x+2 & x^2-x+2 \\ \underline{-x^2-x} & \\ 2x+2 & \\ \underline{2x+2} & \\ 0 & \end{array}$$

Остаток есть 0. Следовательно, общие корни те же, что и у многочленов $x+1$ и 0, то-есть только один общий корень -1.

Мы можем применить то, что сейчас узнали, к другой задаче: узнать, имеет ли заданный многочлен среди своих корней равные, и найти эти равные корни. Мы опять будем предполагать, что имеем дело с многочленом третьей сте-

пени (5). Если этот многочлен имеет своим корнем α, то он делится без остатка на $x-\alpha$. Это вытекает из теоремы Безу, но мы проверим сейчас этот факт простым делением:

$$
\begin{array}{l|l}
x^3+px^2+qx+r & x-\alpha \\
\cline{2-2}
x^3-\alpha x^2 & x^2+(p+\alpha)x+\alpha^2+p\alpha+q \\
\cline{1-1}
(p+\alpha)x^2+qx+r & \\
(p+\alpha)x^2-\alpha(p+\alpha)x & \\
\cline{1-1}
(\alpha^2+p\alpha+q)x+r & \\
(\alpha^2+p\alpha+q)x-\alpha(\alpha^2+p\alpha+q) & \\
\cline{1-1}
\alpha^3+p\alpha^2+q\alpha+r &
\end{array}
$$

Остаток $\alpha^3+p\alpha^2+q\alpha+r$ равен нулю, так как нам дано, что α является корнем уравнения (4). Мы получим заодно, что частное от деления есть

$$x^2+(p+\alpha)x+(\alpha^2+p\alpha+q). \qquad (7)$$

Если бы уравнение (4) имело еще один корень, равный α, то он должен был бы быть корнем многочлена (7). Подставляя α вместо x, получаем $3\alpha^2+2p\alpha+q=0$.

Итак, мы доказали, что равные корни многочлена (5) являются общими корнями этого многочлена и многочлена $3x^2+2px+q$. Этот последний называется *производным многочленом* от многочлена (5). Таким образом, задача привелась к разысканию общих корней многочлена и его производного многочлена, а эту задачу мы уже умеем решать.

Для многочленов n-й степени (2) имеет место вполне аналогичная теорема. Роль многочлена $3x^2+2px+q$ здесь играет многочлен $nax^{n-1}+(n-1)bx^{n-2}+(n-2)cx^{n-3}+\ldots$ $\ldots+k$. Этот многочлен также называется производным от многочлена (2).

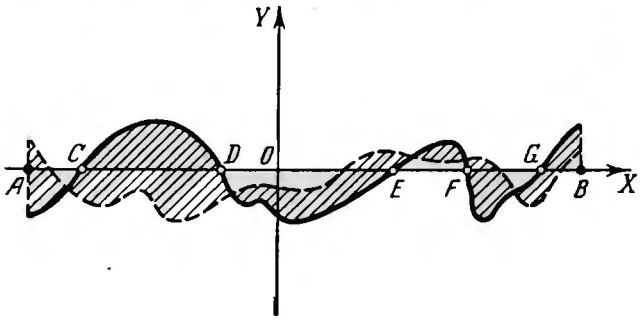
§ 3. ХАРАКТЕРИСТИКА ПАРЫ МНОГОЧЛЕНОВ

Мы переходим теперь к нашей главной задаче. Она формулируется так: задан многочлен f и два числа a и b, причем $a<b$; надо узнать, сколько у многочлена f есть корней, лежащих между a и b,

т. е. меньших b, но бóльших a. Мы увидим дальше, что из решения этой задачи будут следовать ответы на большинство вопросов, которые можно задать о корнях многочленов: об их числе, расположении и даже способе их вычисления.

В решении этой задачи мы будем поступать так же, как и в предшествующем параграфе, т. е. сначала исследуем свойства пары многочленов. При этом мы будем пользоваться двумя методами, которые нам уже встречались: алгебраическим — делением с остатком, и геометрическим — построением графика.

Пусть даны два многочлена, причем нам важен и порядок, в котором они заданы. Первый многочлен обозначим через f_1,



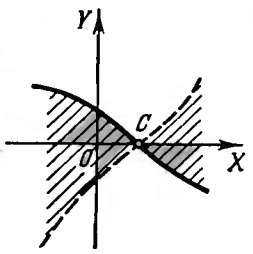
Черт. 3.

а второй — через f_2. Мы будем предполагать, что f_1 и f_2 не имеют общих корней, так как общие их корни мы можем найти по способу предыдущего параграфа. Начертим графики многочленов f_1 и f_2. (График f_1 мы будем чертить жирной линией, а f_2 — пунктиром; черт. 3.) Они будут делить всю плоскость на три части — первую, которая лежит ниже обоих графиков, вторую, которая лежит выше обоих, и третью, которая лежит между двумя графиками. Третья часть и будет нас дальше интересовать. Мы будем ее заштриховывать на чертеже и называть заштрихованной областью. Нанесем на оси x точки с абсциссами a и b и обозначим их через A и B. Будем двигаться от точки A к точке B по оси x. При этом мы будем пересекать графики многочленов f_1 и f_2 в точках, изображающих их корни. В некоторых из этих

точек пересечения мы будем выходить из заштрихованной области, в других — входить в нее. Первые мы будем называть точками выхода, вторые — точками входа. Заметим, что точка не может быть одновременно точкой входа и точкой выхода (как это имеет место на черт. 4 для точки C), так как тогда оно изображала бы общий корень многочленов f_1 и f_2, а мы предположили, что таких нет.

Теперь мы введем одно новое понятие.

Характеристикой многочленов f_1 и f_2 называется разность между числом точек выхода и числом точек входа, принадлежащих графику многочлена f_1 и расположенных между точками A и B.



Черт. 4.

Например, на чертеже 3 точки D и E являются точками выхода, принадлежащими графику f_1, а C, F и G—точками входа на этом же графике. Характеристика в этом случае равна $2-3=-1$. Характеристика является целым числом и обозначается (f_1, f_2). Она, конечно, зависит, кроме многочленов f_1 и f_2, еще и от точек A и B, между которыми мы рассматриваем графики многочленов.

Совершенно очевидно, что характеристика многочленов f_1 и f_2 зависит от того, какой из них мы считаем первым, так как при ее подсчете мы обращаем внимание только на точки входа и выхода на графике этого многочлена. Поэтому интересно выяснить, ка́к именно изменится характеристика, если мы будем считать первым многочленом не f_1, а f_2, или, иначе, ка́к связаны между собой числа (f_1, f_2) и (f_2, f_1).

Удобно представить себе положение вещей на следующей модели. Имеется комната, в которую человек может входить и из которой может выходить. Обозначим число, показывающее, сколько раз он вышел из комнаты за заданный промежуток времени, через P, а сколько раз вошел — через Q. Ясно, что так как за каждым входом, кроме последнего, следует выход, а за выходом — вход, то число входов может отличаться от числа выходов не более чем на 1 в ту или другую сторону. Мы можем это записать так:

$$P = Q + e, \quad e = 0,\ 1 \text{ или } -1. \qquad (8)$$

Более точно значение e определяется следующей таблицей:

1 — если в начальный момент человек был внутри, а в конечный — снаружи;

-1 — если в начальный момент он был снаружи, а в конечный — внутри;

0 — если в начальный и конечный момент времени он был по одну и ту же сторону стен комнаты (черт. 5).

Представим теперь себе, что задача усложнена тем, что в комнате есть две двери — *I* и *II*, и число выходов и входов через каждую дверь подсчитывается отдельно. Число выходов и входов через *I* дверь мы обозначим через P_1 и Q_1, через *II* — P_2 и Q_2. Ясно, что общее число входов и выходов через обе двери подчиняется соотношению (8), т. е.

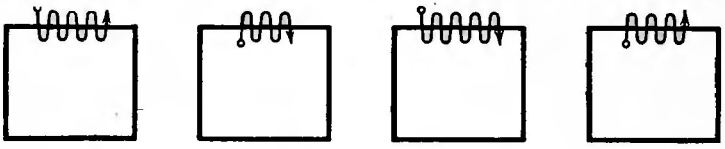
$$P_1 + P_2 = Q_1 + Q_2 + e.$$

Отсюда мы получаем:

$$P_1 - Q_1 = -(P_2 - Q_2) + e.$$

Это показывает, что разности между числом выходов и числом входов в случае *I* и *II* двери просто связаны между собой.

Ясно, какое это имеет отношение к нашим многочленам. Комната — это заштрихованная область, *I* дверь — это график



Черт. 5.

многочлена f_1, *II* дверь — это график f_2. Выходы и входы — это точки входа и точки выхода, $P_1 - Q_1$ — не что иное, как характеристика (f_1, f_2), $P_2 - Q_2$ — характеристика (f_2, f_1). Мы получаем соотношение

$$(f_1, f_2) = -(f_2, f_1) + e,$$

и остается только выяснить, каков смысл e в нашем случае.

Внутри заштрихованной области (комнаты) мы начали или вне ее, зависит от того, принимают ли многочлены f_1 и f_2 при $x = a$ значения одинакового знака или разных знаков. Точно так же, внутри или вне заштрихованной области мы кончили, зависит от того, одинакового или разного знака

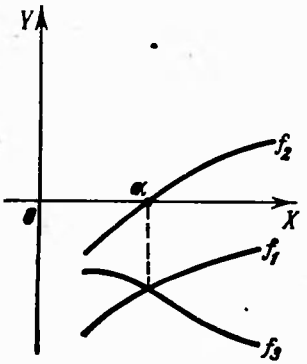
значения f_1 и f_2 при $x = b$. Условимся говорить, что в паре чисел есть одна перемена знака, если они разных знаков, и нет ни одной — если одного, и обозначим значения f_1 и f_2 при $x = a$ через a_1 и a_2, а при $x = b$ через b_1 и b_2. Как легко проверить, e будет совпадать с разностью между числом перемен знака в паре a_1 и a_2 и числом перемен знака в паре b_1 и b_2. Если первое число обозначить через m_1, а второе — через n_1, то $e = m_1 - n_1$.

До сих пор мы не имели способа найти характеристику двух многочленов, так как, для того чтобы воспользоваться определением характеристики, надо было знать их корни. Сейчас мы можем такой способ вывести. Именно здесь вступает в действие деление с остатком. Мы разделим f_1 с остатком на f_2:

$$f_1 = f_2 q + f_3 \quad (f_3 — \text{остаток}), \tag{9}$$

и поставим вопрос, как связаны характеристики (f_1, f_2) и (f_2, f_3). Такая постановка вопроса подсказывается нам результатом предыдущего параграфа, где мы видели, что общие корни у пары многочленов f_1, f_2 те же, что и у пары многочленов f_2, f_3. Прежде всего мы имеем:

$$(f_1, f_2) = -(f_2, f_1) + e. \tag{10}$$



Черт. 6.

Нам остается сравнить (f_2, f_1) и (f_2, f_3). Мы докажем, что эти два числа равны. Для этого заметим, что если α является корнем многочлена f_2, то при $x = \alpha$ значения многочленов f_1 и f_3 совпадают (черт. 6). Это получится, если в (9) заменить x на α и заметить, что член, содержащий f_2, пропадает. Из этого следует, что около точки с абсциссой $x = \alpha$ графики многочленов f_1 и f_3 расположены по одну и ту же сторону от графика многочлена f_2 (черт. 6). Отсюда уже просто следует совпадение характеристик (f_2, f_1) и (f_2, f_3). В самом деле, при подсчете обеих характеристик нам надо перебирать корни многочлена f_2, лежащие между A и B, и смотреть, какие из них являются точками входа, а какие — точками выхода. Но из того, что было только что доказано относительно расположения графиков многочленов f_1, f_2 и f_3, следует, что если корень α является точкой

входа в заштрихованной области, образованной графиками многочленов f_2 и f_1, то он будет изображаться также точкой входа и в области, образованной графиками f_2 и f_3. То же самое имеет место и для точек выхода.

Таким образом, совпадение характеристик (f_2, f_1) и (f_2, f_3) доказано. Вместе с (10) это дает нам соотношение

$$(f_1, f_2) = -(f_2, f_3) + e. \tag{11}$$

Так как многочлены f_2 и f_3 имеют меньшие степени, чем многочлены f_1 и f_2, то это дает нам общий способ вычисления характеристики пары многочленов.

Принципиально наша задача — научиться вычислять характеристику пары многочленов — решена, но можно придать результату более законченный вид. Для этого постараемся избавиться от знака минус в формуле (11). Проще всего взять для этого за f_3 не остаток от деления f_1 на f_2, а этот остаток с обратным знаком. Остаток будет тогда $-f_3$ и (9) будет иметь вид:

$$f_1 = f_2 q - f_3,$$

а формула (11) приобретает более простой вид

$$(f_1, f_2) = (f_2, f_3) + e. \tag{12}$$

Для того чтобы это проверить, нам надо только доказать, что характеристика пары многочленов меняет знак, если второй многочлен меняет знак. Но это совершенно ясно из чертежа, поскольку график второго многочлена заменится тогда симметричной ему относительно оси x кривой, так что все корни первого многочлена, которые были точками выхода, сделаются точками входа и наоборот. Ясно, что характеристика изменит знак.

После того как мы разделим f_1 на f_2 и остаток, взятый с обратным знаком, обозначим через f_3, проделаем то же самое с f_2 и f_3 и будем продолжать, пока у нас не получится остаток, равный нулю. Полученный таким образом ряд многочленов $f_1, f_2, f_3, \ldots, f_k$ (не считая нуля) называется *рядом Штурма* пары многочленов f_1 и f_2. Последний многочлен f_k ряда Штурма не может иметь корней, так как многочлены $f_1, f_2, \ldots, f_k$ только знаками отличаются от тех, которые мы строили в предыдущем параграфе для определения общих корней f_1 и f_2, и если бы f_k имел корень, то он был бы общим корнем f_1 и f_2, а мы предположили, что таких нет. Выписывая формулу (12) для каждой после-

довательной пары многочленов ряда Штурма, мы получим целый ряд формул:

$$(f_1, f_2) = (f_2, f_3) + e_1,$$
$$(f_2, f_3) = (f_3, f_4) + e_2,$$
$$\dots\dots\dots\dots\dots\dots$$
$$(f_{k-2}, f_{k-1}) = (f_{k-1}, f_k) + e_{k-2},$$
$$(f_{k-1}, f_k) = (f_k, 0) + e_{k-1}.$$

Здесь числа $e_1, \dots, e_{k-1}$ имеют прежний смысл, то-есть каждое из них (например, e_i) есть разность между числом перемен знака в паре чисел, являющихся значениями двух рядом стоящих многочленов f_i и f_{i+1} при $x = a$ и при $x = b$.

Заметим теперь, что раз f_k не имеет корней, то характеристика его с любым многочленом равна нулю, в частности $(f_k, 0) = 0$, и мы можем последнюю формулу переписать так: $(f_{k-1}, f_k) = e_{k-1}$. Из предпоследней формулы определим (f_{k-2}, f_{k-1}); оно окажется равным $e_{k-2} + e_{k-1}$.

Поступая так и дальше, мы получим, наконец, что

$$(f_1, f_2) = e_1 + e_2 + \dots + e_{k-1}. \qquad (13)$$

Эту формулу можно еще упростить. Обозначим через $a_1, a_2, \dots, a_k$ значения многочленов $f_1, f_2, \dots, f_k$ при $x = a$, а при $x = b$ — через $b_1, b_2, \dots, b_k$. Если число перемен знаков в паре чисел a_i и a_{i+1} есть m_i, а в паре чисел b_i и b_{i+1} — есть n_i, то

$$e_i = m_i - n_i.$$

Формулу (13) можно поэтому переписать так:

$$(f_1, f_2) = (m_1 - n_1) + (m_2 - n_2) + \dots + (m_{k-1} - n_{k-1}),$$

или

$$(f_1, f_2) = (m_1 + m_2 + \dots + m_{k-1}) - (n_1 + n_2 + \dots + n_{k-1}).$$

Число $m_1 + \dots + m_{k-1}$ показывает, сколько раз в ряду чисел $a_1, a_2, \dots, a_k$ стоят рядом числа противоположных знаков. Число $n_1 + \dots + n_{k-1}$ показывает то же самое для ряда чисел $b_1, b_2, \dots, b_k$. Эти числа называются *числами перемен знаков* в ряду $a_1, \dots, a_k$ и ряду $b_1, \dots, b_k$.

Теперь мы можем сформулировать окончательно полученный нами результат так: *характеристика пары много-*

членов равна разности между числом перемен знаков в значениях многочленов ряда Штурма при $x = a$ *и* $x = b$.

Закончим этот длинный параграф примером. Определим характеристику многочленов $f_1 = x^3 - 8x^2 + 19x - 12$ и $f_2 = x^3 - 9x^2 + 27x - 26$, причем положим $a = 0$, $b = 5$. Найдем ряд Штурма. Делим f_1 с остатком на f_2:

$$\begin{array}{l|l} x^3 - 8x^2 + 19x - 12 & x^3 - 9x^2 + 27x - 26 \\ \underline{x^3 - 9x^2 + 27x - 26} & \underline{\quad 1 \quad} \\ x^2 - 8x + 14 & \end{array}$$

Остаток есть $x^2 - 8x + 14$ и, следовательно, $f_3 = -x^2 + 8x - 14$. Делим f_2 с остатком на f_3:

$$\begin{array}{l|l} x^3 - 9x^2 + 27x - 26 & -x^2 + 8x - 14 \\ \underline{x^3 - 8x^2 + 14x} & \underline{\quad -x + 1 \quad} \\ -x^2 + 13x - 26 & \\ \underline{-x^2 + 8x - 14} & \\ 5x - 12 & \end{array}$$

Остаток есть $5x - 12$ и, следовательно, $f_4 = -5x + 12$. Делим f_3 с остатком на f_4:

$$\begin{array}{l|l} -x^2 + 8x - 14 & -5x + 12 \\ \underline{-x^2 + \frac{12}{5}x} & \underline{\quad \frac{1}{5}x - \frac{28}{5} \quad} \\ \frac{28}{5}x - 14 & \\ \underline{\frac{28}{5}x - 12 \cdot \frac{28}{25}} & \\ -14 + 12 \cdot \frac{28}{25} = -\frac{14}{25} & \end{array}$$

Остаток есть $-\frac{14}{25}$ и, следовательно, $f_5 = \frac{14}{25}$.

Ясно, что следующий остаток уже будет нулем. Ряд Штурма состоит, таким образом, из многочленов $x^3 - 8x^2 + 19x - 12$, $x^3 - 9x^2 + 27x - 26$, $-x^2 + 8x - 14$, $-5x + 12$, $\frac{14}{25}$.

Значения многочленов ряда Штурма выпишем в виде таблицы

	$x=0$	$x=5$
f_1	-12	8
f_2	-26	49
f_3	-14	1
f_4	12	-13
f_5	$\frac{14}{25}$	$\frac{14}{25}$

Таким образом, первый столбец содержит как раз числа $a_1, \ldots, a_k$, а второй — числа $b_1, \ldots, b_k$. В первом столбце есть одна перемена знака, так как рядом стоят -14 и 12, во втором — две, так как рядом стоят 1 и -13 и -13 и $\frac{14}{25}$. Характеристика равна $1-2=-1$.

§ 4. ЧИСЛО КОРНЕЙ МНОГОЧЛЕНА, ЛЕЖАЩИХ МЕЖДУ a и b

Сейчас мы покажем, как от характеристики пары многочленов перейти к нахождению корней одного многочлена. Характеристика пары многочленов f_1 и f_2 равна разности между числом тех корней многочлена f_1 между a и b, которые являются точками выхода, и тех, которые являются точками входа. Если бы мы сумели подобрать многочлен f_2 так, чтобы все корни многочлена f_1 стали точками выхода и, значит, ни один не был бы точкой входа, то характеристика пары многочленов f_1 и f_2 в точности совпала бы с числом корней многочлена f_1, лежащих между a и b.

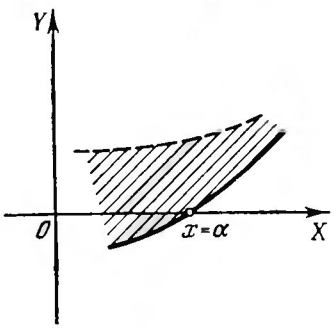
Посмотрим, как подбирать такой многочлен f_2. При этом нам нужно будет предполагать, что многочлен f_1 не имеет равных корней. Пусть α есть корень многочлена f_1. Как должен быть расположен график многочлена f_2, чтобы точка, изображающая корень α, была точкой выхода? Тут возможны два случая в зависимости от того, поднимается ли график многочлена f_1 при $x=\alpha$ снизу вверх или спускается

сверху вниз. Как показывают черт. 7 и 8, в первом случае график многочлена f_2 должен лежать при $x = \alpha$ выше, а во втором — ниже оси x. Иными словами, в первом случае значение многочлена f_2 при $x = \alpha$ должно быть положительным, а во втором — отрицательным. Посмотрим теперь, как научиться узнавать, какой из двух случаев имеет место. Первый случай заключается в том, что при x, меньшем α, многочлен f_1 отрицателен, а при x, большем α, — положителен. Во втором же случае — наоборот.

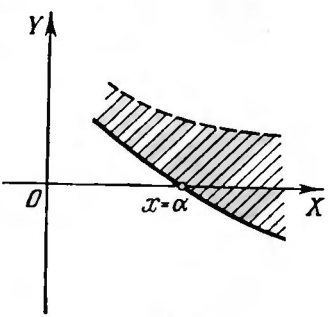
Предположим опять, что мы имеем дело с многочленом третьей степени $f_1 = x^3 + px^2 + qx + r$, имеющим корень α, и вспомним, что еще в § 2 мы показали, что такой многочлен делится на $x - \alpha$ и частное равно $x^2 + (p + \alpha)x + (\alpha^2 + p\alpha + q)$:

$$x^3 + px^2 + qx + r = (x - \alpha)[x^2 + (p + \alpha)x + (\alpha^2 + p\alpha + q)].$$

Воспользуемся этим разложением, чтобы определить, каковы знаки тех значений, которые принимает многочлен f_1,



Черт. 7.



Черт. 8.

когда x немного меньше или немного больше α. Относительно первого множителя все ясно: если x меньше α, то он отрицателен, а если больше — то положителен. Обратимся ко второму множителю. При $x = \alpha$ он не обращается в нуль. В самом деле, мы видели в § 2, что это свидетельствовало бы о наличии у многочлена f_1 двух равных корней, а мы предположили, что их нет. Значение второго множителя при $x = \alpha$ равно как раз значению производного многочлена f_1 при $x = \alpha$. Если это — положительное число, то и при x, не слишком сильно отличающемся от α (именно, если x ближе

к α, чем какой-либо из корней этого квадратного трехчлена), второй множитель положителен, а это значит, что все произведение ведет себя в смысле знаков принимаемых значений как первый множитель; иными словами, мы имеем здесь первый случай. Наоборот, если значение производного многочлена отрицательно, то все знаки меняются на обратные, и мы имеем второй случай.

Мы установили, что здесь решающую роль играет уже знакомый нам производный многочлен, а именно, первый случай имеет место, если значение производного многочлена при $x = \alpha$ положительно, и второй — если отрицательно. То же самое имеет место и для многочленов любой степени.

Но из этого мы видим сразу, что производный многочлен как раз и является таким многочленом f_2, какой нам надо найти. Иными словами, если за f_2 взять производный многочлен многочлена f_1, то характеристика (f_1, f_2) будет равна числу корней многочлена f_1, расположенных между a и b. Так как мы вывели в предыдущем параграфе способ вычисления характеристики, то мы имеем способ нахождения числа корней любого многочлена, расположенных между a и b.

Сформулируем полученное правило для отыскания числа корней многочлена.

Для того чтобы найти число корней многочлена, расположенных между a и b, надо найти ряд Штурма для многочлена и его производного многочлена и вычислить значения многочленов этого ряда сначала при $x = a$, потом при $x = b$. Разность между числом перемен знаков в первом и втором рядах чисел и будет искомым числом корней.

Это правило является знаменитой теоремой французского математика Штурма. Она доказана им в 1829 г.

Интересно, что рассуждения, при помощи которых мы решили нашу задачу — найти число корней многочлена, лежащих между a и b, — имеет много общего с теми рассуждениями, при помощи которых мы в § 2 решили другую задачу — найти равные корни многочлена. В обоих случаях оказалось, что проще сначала исследовать некоторое свойство пары многочленов — в первом случае это был вопрос об общих корнях, во втором — о нахождении характеристики. Причина того, что в свойствах пары многочленов легче разобраться, чем в свойствах одного многочлена, в обоих случаях была одна и та же — к двум многочленам можно

применять деление с остатком, благодаря чему решение интересующего нас вопроса о двух данных многочленах сводится к решению того же вопроса о двух многочленах меньшей степени. Затем мы в обоих случаях подбирали к заданному нам многочлену f_1 такой вспомогательный многочлен f_2 (в обоих случаях за такой многочлен можно принять производный многочлен к f_1), что решенный нами вопрос о паре многочленов в применении к паре f_1, f_2 дает решение интересующего нас вопроса о многочлене f_1.

Покажем теперь, как при помощи теоремы Штурма можно решить ряд задач о корнях многочленов любой степени.

Прежде всего мы обратимся к задаче об определении числа всех, а не только лежащих между a и b, корней заданного многочлена f. Решение этой задачи получается сразу же, если соединить теорему Штурма с результатами, полученными нами в § 1. Там мы показали, как найти такое число N, что все корни многочлена лежат между $-N$ и N. Из этого следует, что если мы при помощи теоремы Штурма определим число корней многочлена, расположенных между $-N$ и N (взяв $a=-N$, $b=N$), то получим число всех его корней. Таким образом, задача решена.

Это решение можно, однако, представить в более изящной форме. Пусть ряд Штурма для многочлена f_1 и его производного многочлена f_2 состоит из многочленов $f_1, f_2, \ldots, f_k$.

Для каждого многочлена f_i из этого ряда существует согласно § 1 такое свое число N_i, что корни f_i лежат между $-N_i$ и N_i. Выберем теперь за N любое число, большее всех чисел N_i. Тогда корни всех многочленов f_i расположены между $-N$ и N. Если мы теперь применим теорему Штурма, взяв $a=-N$ и $b=N$, то мы получим, конечно, число всех корней нашего многочлена. Но оказывается, что теперь при вычислении числа перемен знаков в значениях многочленов рядов Штурма при $x=a$ и при $x=b$ нам нет надобности вычислять сами эти значения.

В самом деле, нас ведь интересуют только знаки, а мы показали в § 1, что знак того значения, которое принимает любой многочлен при x, меньшем $-N$ или большем N, совпадает со знаком его первого члена. Мы же так выбрали N, что оно больше всех N_i, а значит, $-N$ меньше всех N_i. Таким образом, знаки значений многочленов ряда Штурма при $x=N$ и при $x=-N$ определяются знаками их первых членов. Отметим тот любопытный факт, что, хотя мы

пользовались в рассуждениях существованием числа N, нам нет необходимости вычислять его в каждом конкретном случае, так как первый член каждого многочлена ряда Штурма имеет вид ax^k и его знак при $x=-N$ или $x=N$ вовсе не зависит от величины N.

Пример. Найти число всех корней многочлена $f_1 = x^3+3x-1$.

Производный многочлен есть $3x^2+3=f_2$. Находим ряд Штурма. Делим f_1 с остатком на f_2:

$$\begin{array}{l|l} x^3+3x-1 & 3x^2+3 \\ \underline{x^3+x} & \overline{1/3x} \\ 2x-1 & \end{array}$$

Остаток есть $2x-1$, следовательно, $f_3=-2x+1$. Делим f_2 с остатком на f_3:

$$\begin{array}{l|l} 3x^2+3 & -2x+1 \\ \underline{3x^2-3/2x} & \overline{-3/2x-3/4} \\ 3/2x+3 & \\ \underline{3/2x-3/4} & \\ 3\,3/4 & \end{array}$$

Остаток есть $3\,3/4$, следовательно, $f_4=-3\,3/4$. Ясно, что следующий многочлен уже равен нулю. Ряд Штурма состоит, таким образом, из многочленов x^3+3x-1, $3x^2+3$, $-2x+1$ и $-3\,3/4$.

Если мы теперь выберем число N (конечно, положительное) так, как было указано, то знаки значений этих многочленов при $x=-N$ и $x=N$ будут совпадать со знаками их первых членов. Эти знаки даны в таблице

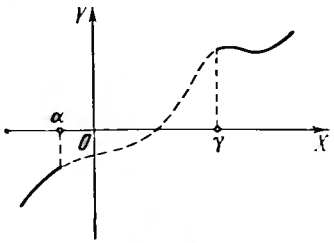
	$x=-N$	$x=N$
f_1	$-$	$+$
f_2	$+$	$+$
f_3	$+$	$-$
f_4	$-$	$-$

В первом столбце имеются две перемены знака, во втором — одна. Многочлен x^3+3x-1 имеет, таким образом, $2-1=1$ корень.

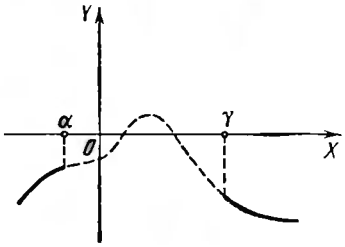
Теперь мы можем перейти к задаче о вычислении корней многочлена. Мы поставим себе целью вычислить все корни с какой угодно степенью точности. Собственно, ничего лучшего мы не имеем и в случае квадратных уравнений, хотя и кажется, что там формула дает точное, а не приближенное значение корня. Действительно, если мы говорим, что корень многочлена $x^2 - a$ равен $\sqrt{a}$, то этим мы только указываем, что если применить правило приближенного вычисления квадратных корней, то получится приближенное значение корня многочлена с любой степенью точности.

Для приближенного вычисления корней многочленов нам надо прежде всего (в отличие от того, что имело место при определении числа корней) действительно вычислить по правилу § 1 такое число N, что между $-N$ и N лежат все корни нашего многочлена. После этого мы делим отрезок между $-N$ и N на части, например пополам, и находим по теореме Штурма число корней в каждой из половин, то-есть между $-N$ и 0 и между 0 и N. С каждым из этих отрезков, если оказалось, что в нем содержатся корни, поступаем так же. Так мы делаем до тех пор, пока не раздробим наш отрезок от $-N$ до N на отрезки сколь угодно малой длины, причем нам известно, в каких из этих маленьких отрезков лежат корни, а в каких — нет. Но это и значит, что мы смогли вычислить корни с какой угодно степенью точности, так как если мы, например, узнали, что корень лежит между числами α и β, разность между которыми меньше 0,001, то α и является приближенным значением корня с недостатком с точностью до 0,001, а β — значением с избытком. Для удобства вычислений часто делят отрезок не пополам, а на 10 частей. Когда мы дойдем до таких маленьких интервалов, что в каждом из них лежит или один корень многочлена, или ни одного, то при следующих вычислениях нам нет надобности пользоваться теоремой Штурма и вычислять значения многочленов ряда Штурма. В самом деле, пусть мы знаем, что между α и β лежит ровно один корень многочлена. Пусть мы разделили числом γ отрезок между α и β на две части и хотим узнать, в какой из этих частей, между α и γ или между γ и β, лежит корень. Тогда достаточно вычислить значения многочлена при $x = \alpha$ и при $x = \gamma$. Если они разных знаков, то, как видно из черт. 9, при $x = \alpha$ и при $x = \gamma$ график многочлена лежит по разные стороны оси x, а следовательно, должен где-то между α

и γ эту ось пересечь. Но это и значит, что корень лежит между α и γ. Наоборот, если значения многочлена при $x=\alpha$ и при $x=\gamma$ одного знака, то корень не может лежать между α и γ. Действительно, тогда график многочлена где-то между α и γ переходил бы с одной стороны оси x на другую, а значит, где-то в другом месте, тоже между α и γ, должен был бы вернуться назад, раз при $x=\alpha$ и при $x=\gamma$ он лежит по одну сторону оси x (черт. 10). Но это означало бы, что многочлен имеет между α и γ по крайней мере два корня, а этого не может быть, так как мы предположили, что между α и β есть всего один корень. Значит, корень не может лежать между α и γ, а лежит между γ и β.



Черт. 9.



Черт. 10.

Пример. Вычислим с точностью до 0,1 корень многочлена x^3+3x-1. В прошлом примере мы видели, что этот многочлен имеет один единственный корень, значит, мы можем применить то упрощение в общем методе, о котором только что говорили. Вычислим прежде всего N. Согласно правилу § 1 нам надо за N взять наибольшее из чисел

$$0,\ \sqrt{3\cdot 3}\ \text{ и }\ \sqrt[3]{3}.$$

Наибольшее из них, конечно, $\sqrt{3\cdot 3}=3$ и, значит, $N=3$. Таким образом, искомый корень лежит между -3 и 3.

Узнаем теперь знак корня. Для этого найдем значения многочлена при $x=0$ и при $x=3$. Получим -1 и 36. Так как это числа разных знаков, то корень лежит между 0 и 3, а следовательно, положителен.

Найдем теперь значения многочлена при $x=1$ и $x=2$. Получим 3 и 13. Таким образом, значения при $x=1$, 2 и 3 — все одного знака, а значения при $x=0$ и при $x=1$ —

рааных. Это значит, что корень лежит между 0 и 1, т. е. $x = 0, \ldots$ Чтобы найти первый знак после запятой, надо узнать, в каком из десяти промежутков (между 0 и $^1/_{10}$, между $^1/_{10}$ и $^2/_{10}$, ..., между $^9/_{10}$ и 1) лежит корень. Для этого положим сначала $x = {}^5/_{10} = {}^1/_2$. Получим в качестве значения многочлена $^5/_8$. Так как -1 и $^5/_8$ разных знаков, то корень лежит между 0 и $^1/_2$. Положим теперь $^3/_{10}$. Получим значение, равное $\frac{27}{1000} - \frac{9}{10} + 1 = \frac{27}{1000} - \frac{1}{10}$. Это отрицательное число, и, значит, корень лежит между $^3/_{10}$ и $^5/_{10}$, так как значения многочлена при $x = {}^3/_{10}$ и $x = {}^5/_{10}$ — разных знаков. Нам остается положить $x = {}^4/_{10}$. Получим значение $\frac{64}{1000} + \frac{12}{10} - 1 = \frac{64}{1000} + \frac{2}{10}$. Так как это положительное число, то корень лежит между $^3/_{10}$ и $^4/_{10}$. Наша задача решена. Мы нашли, что с точностью до 0,1 значение корня многочлена $x^3 + 3x - 1$ есть 0,3.